SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Knockdown and cDNA expression

Snail knockdowns were performed with lentiviral miR-30 based shRNA (Snai1.1197) cloned into pTRIPZ and a pLKO shRNA targeting Snail (designated shRNA 4, target sequence CCACTCGGATGTGAAGAGATA). All other knockdowns were performed with mir-E based shRNAs cloned into lentiviral vector LT3GEPIR. Tandem shRNAs were cloned as described (Chicas et al., 2010) into a lentiviral tandem vector. MiR-30 based shRNAs were cloned using 97-mer oligos as described (Fellmann et al., 2013). Sequences of 97mer oligos used for cloning are listed below.

Snai1.1197	TGCTGTTGACAGTGAGCGCGAGGTACAACAGACTATGCAATAGTGAAGCCACAGATGTATTGCATAGTCTGTTGTACCTCATGCCTACTGCCTCGGA
Zeb1.1036	TGCTGTTGACAGTGAGCGCCCTGTGGATTATGAGTTCAAATAGTGAAGCCACAGATGTATTTGAACTCATAATCCACAGGTTGCCTACTGCCTCGGA
Zeb1.2505	TGCTGTTGACAGTGAGCGATGGCATATACATATTCAGCTATAGTGAAGCCACAGATGTATAGCTGAATATGTATATGCCACTGCCTACTGCCTCGGA
Sox4.965	TGCTGTTGACAGTGAGCGCCAGCGACAAGATTCCGTTCATTAGTGAAGCCACAGATGTAATGAACGGAATCTTGTCGCTGTTGCCTACTGCCTCGGA
Sox4.2137	TGCTGTTGACAGTGAGCGCTAGATGGAGAGTAGAAGGAGATAGTGAAGCCACAGATGTATCTCCTTCTACTCTCCATCTATTGCCTACTGCCTCGGA
Sox4.2091	TGCTGTTGACAGTGAGCGCAAAGAAGAAGAAAGAAAGAAATAGTGAAGCCACAGATGTATTTCTTTCTTTCTTCTTCTTTTTGCCTACTGCCTCGGA
Renilla.713	TGCTGTTGACAGTGAGCGCAGGAATTATAATGCTTATCTATAGTGAAGCCACAGATGTATAGATAAGCATTATAATTCCTATGCCTACTGCCTCGGA
Klf5.1343	TGCTGTTGACAGTGAGCGCCAGATACAACAGAAGGAGTAATAGTGAAGCCACAGATGTATTACTCCTTCTGTTGTATCTGATGCCTACTGCCTCGGA
Klf5.636	TGCTGTTGACAGTGAGCGACAGTTCTTCACTGACACTGAATAGTGAAGCCACAGATGTATTCAGTGTCAGTGAAGAACTGGTGCCTACTGCCTCGGA
Klf5.1704	TGCTGTTGACAGTGAGCGCACCGCAGACCTAACTTCATAATAGTGAAGCCACAGATGTATTATGAAGTTAGGTCTGCGGTTTGCCTACTGCCTCGGA
Foxa2.1649	TGCTGTTGACAGTGAGCGACACAGTGATCTGTCATTCTAATAGTGAAGCCACAGATGTATTAGAATGACAGATCACTGTGGTGCCTACTGCCTCGGA

For Smad4 reintroduction, human SMAD4 was cloned into the pLVX-Tight-Puro (TetOn) or pLVX-IRES-Hyg (constitutive) vectors. For overexpression, cDNAs corresponding to Snail, Sox4, and Bcl-X_L were cloned into pLVX-Tight-Puro. TetON inductions were done at 1 μg/ml of doxycycline.

In vitro cell-based assays

For in vitro cleaved Caspase activity measurements, cells were plated in 96-well format at 500 cells/well for 12-24 h prior to treatment with MK2206 (Tocris, 2.5 μM), and SB505124 (Sigma, 2.5 μM) or TGF-β (R&D; 100 pM). Cells were assayed for cleaved Caspase 3/7 activity using CaspaseGlo (Promega) and normalized to cell titer using Cell TiterGlo (Promega). Experiments were performed in triplicate, and p-values were calculated using Student's unpaired t-test. For oncosphere formation assays, 500 cells/well

were plated in ultra-low attachment 96 well plates (Corning) in DMEM/F-12 medium supplemented with B-27 serum replacement and 5 μg/ml Heparin. Organoids were cultured as described (Boj et al., 2015).

Immunoblotting

Antibodies against Snail (#3879), Zeb1 (#3396), E-cadherin (#3195), Smad2/3 (#3102), and cleaved Caspase 3 (#9661) were purchased from Cell Signaling. Klf5 antibodies were from Abcam (ab137676). Tubulin and Gapdh antibodies were purchased from Sigma. The Sox4 antibody was purchased from Diagenode (C15310129). Hnf4a (H-171) and Sox17 (S-20) antibodies were purchased from Santa Cruz. Pdx1 (ab47267) and Cdx2 (ab15258) antibodies were purchased from Abcam. Klf4 antibody was purchased from ProSci (PM-6141). Foxa2 antibody was purchased from Millipore (17-10258).

IF and IHC

IF and IHC were performed on tissues fixed with 4% paraformaldehyde overnight, followed by embedding and sectioning. Antigen retrieval was performed using citrate buffer. IHC was performed using Impress Reagent and DAB substrate purchased from Vector labs. IF secondary antibodies were purchased from Life Technologies. For Sox4 IHC, antibodies from Novus were used (NBP1-89506). IF antibodies for E-cadherin were from Cell Signaling (#3195) and Klf5 antibodies were from Sigma (SAB4200338). Antibodies derived from the Troma III hybridoma were used to detect CK19. GFP antibodies were from Aves Labs (GFP-1020). Cleaved Caspase 3 antibodies were from Cell Signaling (#9661).

TCGA data

All TCGA data was accessed through cBioportal (Gao et al., 2013).

ChIP and ChIP-seq

ChIP-seq experiments were performed with 30-50 million PDA cells each. ChIP-seq was performed using anti-Klf5 (Abcam, ab137676), anti-Sox4 (Diagenode, C15310129), Smad2/3 (CST, #3102), and Med1 (Bethyl, A300-793A). Briefly, cells were crosslinked at room temperature for 10 min with 1% formaldehyde, washed twice with PBS and sonicated in lysis buffer (50 mM HEPES/KOH pH 7.5, 140 mM NaCl, 0.1% Na-Deoxycholate, 1% Triton X-100, 1 mM EDTA, complete protease inhibitor cocktail). Samples were immunoprecipitated at 4 ºC overnight and washed six times with ChIP wash buffer (20 mM Tris, pH 7.9, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS) and once with TE buffer. Bound DNA was eluted in 1% SDS buffer and reverse-crosslinked for 6 h at 65 ºC. DNA samples were treated sequentially with RNase A and Protease K and then purified with ChIP DNA Clean & Concentrator (Zymo Research). ChIP DNA was subjected to qPCR analysis or ChIP-seq library generation. For Smad2/3 ChIP, primer sequences were as follows: *Klf5* locus (F- GCGCTCTTGGGTACTGACTC, R-AAGGGTAGGAAGCAGGAAGC), *Cdh1* locus (F- GCCTTAACAGCCGTTCTCAG, R-TTTTCCTGAGCCAAGTCCAC), *Serpine1* locus (F- GCTTCCATATCTGGGTCAGG, R-CTGCCACCTCAGCTTCTCTT), *Smad7* locus (F- CTGTCCTGTCTGTCCATCCA, R-CCCAAATGGCAGTTTGAAGT). ChIP-Seq DNA samples were quantified and quality assessed by Ribogreen and Agilent Bioanalyzer. DNA fragments range from 200-600bp were selected and constructed for ChIP-Seq library with TruSeq ChIP Sample Prep Kit (Illumina) according to manufacturer's instructions. Sequencing libraries were multiplexed and ran on a Hiseq 2500 platform. Single end (50bp) FASTQ reads were mapped to mouse genome mm10 (GRCm38, Dec/2011) with Bowtie2 with default filtering criteria for single end reads (Langmead and Salzberg, 2012). Resulted SAM files were converted to BAM files though "*view*" function in Samtools 0.1.19 (Li et al., 2009). BAM files were sorted and indexed with "*sort*" then "*index*" function in Samtools (Li et al., 2009). To visualize ChIP-Seq data, BAM files were converted to TDF file by IGV Tools 2.3.32 (Robinson et al., 2011) using the command "*igvtools count –z 5 –w 25 –e 250*", specifying the coverage window size to be 25bp and average fragment size of 250bp. TDF files are loaded into IGV genome browser (Robinson et al., 2011) and signal intensities are normalized by "$1x10^6$/total million reads" or Reads Per Million Reads (RPM) to display normalized coverage data tracks. Gene track views were exported and visualized with Adobe Illustrator. Scale bar for length of genomic ranges was indicated above each gene track plot. Peak calling from ChIP-Seq data was performed with MACS 1.4.2(Zhang et al., 2008) and verified by HOMER (v4.2) (Heinz et al., 2010). The parameters for peak calling were p value < 1e-8 to detect high confidence binding events with elevated stringency. Input samples were used as reference controls for background correction. Peaks identified from MACS 1.4.2 are annotated with HOMER (v4.2) using *annotatePeaks.pl* function. Genes are assigned with

the “nearest TSS” criteria. Peak region overlap is performed with the *intersect* function from Bedtools 2.17.0 (Quinlan and Hall, 2010). Tag density for genomic ranges surrounding defined peak centers were calculated using *annotatePeaks.pl* function in HOMER (v4.2)(Heinz et al., 2010). Log2 transformed tag densities were pre-ranked by peak score. Data matrix from each ChIP-Seq experiment were merged by peak names and plotted for heatmaps in *R* (v3.0.1). White indicates low tag density and blue indicates high tag density consistently. Motif analysis was performed using P-Scan (Zambelli et al., 2009).

References

Boj, S.F., Hwang, C.I., Baker, L.A., Chio, II, Engle, D.D., Corbo, V., Jager, M., Ponz-Sarvise, M., Tiriac, H., Spector, M.S.*, et al.* (2015). Organoid models of human and mouse ductal pancreatic cancer. Cell *160*, 324-338.

Chicas, A., Wang, X., Zhang, C., McCurrach, M., Zhao, Z., Mert, O., Dickins, R.A., Narita, M., Zhang, M., and Lowe, S.W. (2010). Dissecting the unique role of the retinoblastoma tumor suppressor during cellular senescence. Cancer cell *17*, 376-387.

Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., Lai, D.Y., Barbosa, I.A., Kwon, J.S., Guan, Y.*, et al.* (2013). An optimized microRNA backbone for effective single-copy RNAi. Cell reports *5*, 1704-1713.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E.*, et al.* (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science signaling *6*, pl1.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature biotechnology *29*, 24-26.

Zambelli, F., Pesole, G., and Pavesi, G. (2009). Pscan: finding over-represented transcription factor binding site motifs in sequences from co-regulated or co-expressed genes. Nucleic acids research *37*, W247-252.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.